Statistics: Lecture 3 - (Continuous) Probability Distributions and Bivariate Extensions

Jan Bauer

jan.bauer@dhbw-mannheim.de

25.11.19

(Soong 3.3, 3.4, 4., 4.1, 7-7.4, 9.3.2.2.)

Table of Contents

1. Continuous Probability Distributions
2. Bivariate Random Variables
3. Bivariate Distribution: Independence, (Conditional) Moments

Table of Contents

1 Continuous Probability Distributions

2 Bivariate Random Variables

3 Bivariate Distribution: Independence, (Conditional) Moments

Uniform Distribution

A continuous rv X has a uniform distribution over an interval $[a,b]$ if it is equally likely to take on any value in this interval. We write $X \sim U(a,b)$ and it holds

- pdf:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a,b] \\ 0 & \text{else} \end{cases}.$$

- CDF:

$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & x \in [a,b] \\ 1 & x > b \end{cases}.$$

- Moments:

$$\mathrm{E}(X) = \frac{a+b}{2} \quad \text{and} \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

DHBW Duale Hochschule Baden-Württemberg Mannheim

Uniform Distribution Example

Let $X \sim U(a,b)$. Show that $\mathrm{E}(X) = \frac{a+b}{2}$:

We recap that $\int x \, \mathrm{d}x = \frac{1}{2}x^2$. Then $\mathrm{E}(X) \ldots$

$$
\begin{aligned}
&\equiv \int_{-\infty}^{\infty} x \cdot f_X(x) \, \mathrm{d}x \\
&\equiv \int_{-\infty}^{a} x \cdot f_X(x) \, \mathrm{d}x + \int_{a}^{b} x \cdot f_X(x) \, \mathrm{d}x + \int_{b}^{\infty} x \cdot f_X(x) \, \mathrm{d}x \\
&\equiv \int_{-\infty}^{a} 0 \, \mathrm{d}x + \int_{a}^{b} x \frac{1}{b-a} \, \mathrm{d}x + \int_{b}^{\infty} 0 \, \mathrm{d}x \\
&= \frac{1}{b-a} \int_{a}^{b} x \, \mathrm{d}x \\
&= \frac{1}{b-a} \left(\frac{1}{2} b^2 - \frac{1}{2} a^2 \right) \\
&= \frac{b^2-a^2}{2(b-a)} \\
&= \frac{b^2-a^2}{2(b-a)} \\
&= \frac{(b-a)(b+a)}{2(b-a)} \\
&= \frac{b+a}{2}
\end{aligned}
$$

Gaussian Distribution

A continuous rv X is Gaussian (or normal distributed) if its pdf is of the following form. We write $X \sim N(\mu, \sigma^2)$.

- pdf:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), x \in (-\infty, \infty), \sigma > 0 .$$

- CDF:

$$F_X(x) \equiv \int_{-\infty}^{x} f_X(u)\ \mathrm{d}u \equiv \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) \mathrm{d}u .$$

- Moments:

$$\mathrm{E}(X) = \mu \quad \text{and} \quad \mathrm{Var}(X) = \sigma^2$$

Gaussian Distribution Density Example

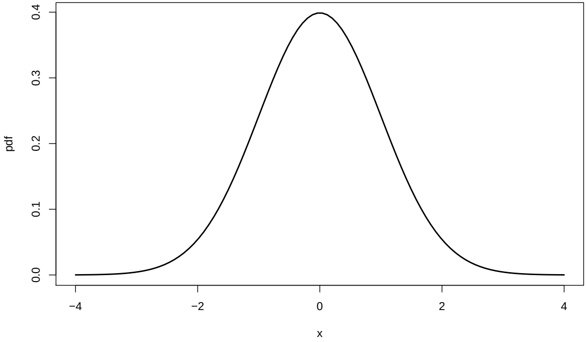

Figure: $N(0, 1)$ Distribution (pdf)

Gaussian Density Example

Recap: The pdf is a function for which the area under the curve corresponding to any interval is equal to the probability that X will take on a value in that interval

$$\rightsquigarrow \mathrm{P}(-1 \leq X) = \int_{-1}^{\infty} f_X(x)\, \mathrm{d}x$$

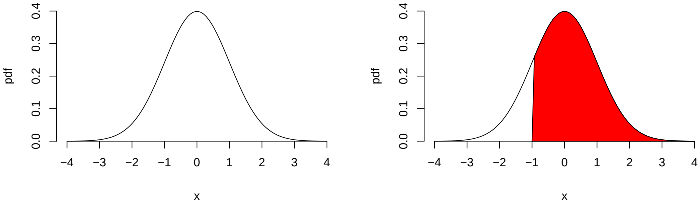

Figure: pdf of $N(0,1)$

Gaussian Distribution Density Example

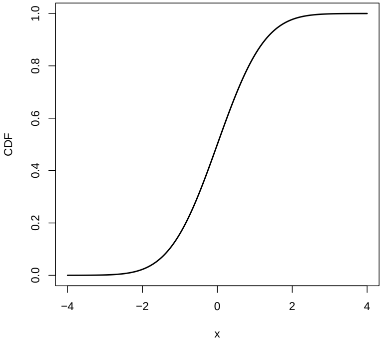

Figure: $N(0,1)$ Distribution (CDF: $F_X(x) = \mathrm{P}(X \leq x)$)

Recap (Example): $\mathrm{P}(X \leq 0.5) = F_X(0.5) = 0.5.$

$$\mathrm{P}(X \leq 5) = F_X(5) \approx 1.$$

Gaussian Density Example

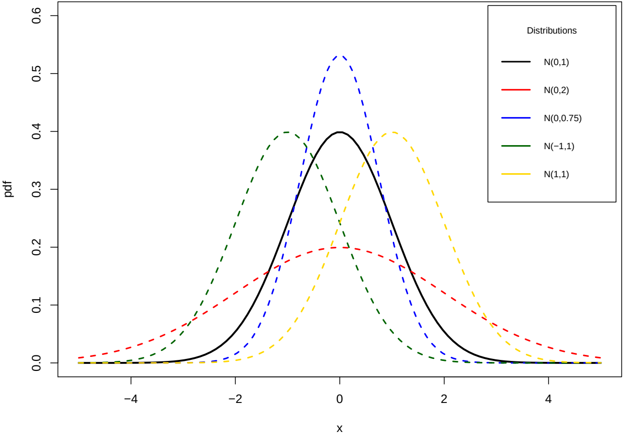

Figure: Comparison of Gaussian Distributions (pdf)

Gaussian Density Interpretation

What does the bell-curve tell us?

- μ: Center of the graph
- σ^2: Height and width of the graph
  - low and wide: Large variance
  - high and narrow: Small variance

Gaussian Distribution Usefulness

- Many variables are Gaussian by nature
- Many distributions convergence against the Gaussian distribution
- Central Limit Theorem: Sum of rvs converges (under some assumptions) against the Gaussian distribution
- Will be useful for "Estimation" and "Hypothesis Tests"

Gaussian Density Example: Height of this class is close to Gaussian

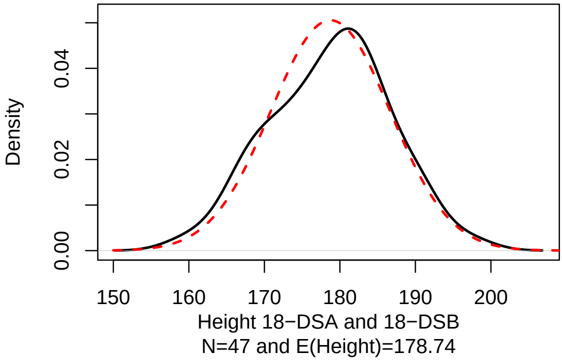

Figure: (Kernel estimated) pdf (-) and Gaussian pdf (-)

χ^2 Distribution

A continuous rv X is χ^2 distributed (with n degrees of freedom) if its pdf is of the following form. We write $X \sim \chi^2$ (or $X \sim \chi^2_n$).

- pdf:

$$f_X(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}} & x \geq 0 \\ 0 & \text{else} \end{cases}$$

- Exptected value:

$$\mathrm{E}(X) = n$$

- Variance:

$$\mathrm{Var}(X) = 2n$$

χ^2 Density Example

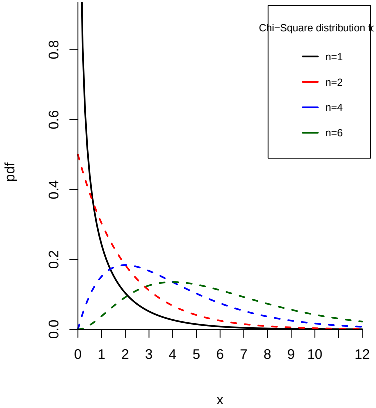

Figure: pdf of χ^2 for df $n = 1, 2, 4, 6$

(Student's) t-Distribution

A continuous rv X is (Student's) t-distributed (with ν degrees of freedom) if its pdf is of the following form. We write $X \sim t$ (or $X \sim t_\nu$).

- pdf:

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

- Expected value:

$$\mathrm{E}(X) = 0 \quad (\nu > 1)$$

- Variance:

$$\mathrm{Var}(X) = \begin{cases} \frac{\nu}{\nu-2} & \nu > 2 \\ \infty & 1 < \nu \leq 2 \end{cases}$$

- Note: It holds that $t_\nu \to N(0,1) \quad (\nu \to \infty)$

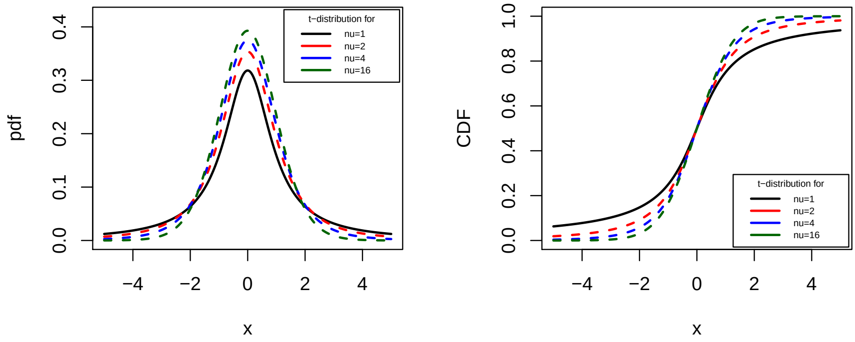

Figure: pdf and CDF of t_ν for df $\nu = 1, 2, 4, 16$

DHBW Duale Hochschule Baden-Württemberg Mannheim

Table of Contents

1 Continuous Probability Distributions

2 Bivariate Random Variables

3 Bivariate Distribution: Independence, (Conditional) Moments

Table of Contents

1 Continuous Probability Distributions

2 Bivariate Random Variables

3 Bivariate Distribution: Independence, (Conditional) Moments

Joint Cumulated Distribution Function

- Motivation: Some outcomes have to be described by two (or more) outcomes simultaneously $\rightsquigarrow$ consider two rvs X and Y.

- Recap: The CDF provides the probability that X will be at or below any given value x. We define it by

$$F_X(x) \equiv \mathrm{P}(X \leq x) .$$

- Analogue: The (joint) CDF provides the probability that (X, Y) will be at or below any given value (x, y). We define it by

$$F_{XY}(x, y) = \mathrm{P}(X \leq x \cap Y \leq y) \quad \forall x, y .$$

Duale Hochschule Baden-Württemberg Mannheim

CDF Properties

- $F_{XY}(x,y)$ exists both for discrete and for continuous rvs

- $F_{XY}(x,y) \in [0,1]$

- For the limits, we have
  - $\lim_{x,y\to-\infty} F_{XY}(x,y) = \lim_{x\to-\infty} F_{XY}(x,y) = \lim_{y\to-\infty} F_{XY}(x,y) = 0$
  - $\lim_{x,y\to\infty} F_{XY}(x,y) = 1$
  - $\lim_{x\to\infty} F_{XY}(x,y) = F_X(x)$
  - $\lim_{y\to\infty} F_{XY}(x,y) = F_Y(y)$

- The CDF is non-decreasing

Discrete Random Variables pmf (Recap)

For now on, we consider a discrete random variable X associated with the distinct outcomes x_i, $i = 1, 2, \ldots$.

- The function

$$f_X(x) \equiv \mathrm{P}(X = x) \quad \forall x$$

  is defined as the probability mass function (pmf) of X.

- What we might have expected already:

$$0 < f_X(x_i) \leq 1 \ \forall i \quad \text{and} \quad \sum_i f_X(x_i) = 1 \, .$$

Note: Those are necessary conditions for a pmf

DHBW Duale Hochschule Baden-Württemberg Mannheim

Discrete Random Variables (joint) pmf

For now on, we consider two discrete random variables X and Y associated with the distinct outcomes (x_i, y_j) $i, j = 1, 2, \ldots$.

- The function

$$f_{XY}(x, y) \equiv \mathrm{P}(X = x, Y = y) \;\; \forall x, y$$

  is defined as the (joint) probability mass function (pmf) of X and Y.

- What we might have expected already:
  - $0 < f_{XY}(x_i, y_j) \leq 1$ (*necessary* condition)
  - $\sum_i \sum_j f_{XY}(x_i, y_j) = 1$ (*necessary* condition)
  - $\sum_i f_{XY}(x_i, y) = f_Y(y)$
  - $\sum_j f_{XY}(x, y_j) = f_X(X)$

  where $f_X(x)$ and $f_Y(y)$ are called marginal probability mass function (marginal pmf) of X and Y, respectively.

Relations between discrete CDF and pmf

From the definition of the CDF, we conclude that

$$F_{XY}(x,y) = \sum_{i}^{i:x_i \leq x} \sum_{j}^{j:x_j \leq x} f_{XY}(x_i, y_j)\,.$$

probability mass function Example

Consider the probability mass function, $f_{XY}(x, y)$, for the discrete random variables X and Y:

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to \frac{xy - 1}{6}$$

where $\Omega = \{1, 2\} \times \{1, 2\}$.

1. Show that $\sum_i \sum_j f_{XY}(x_i, y_j) = 1$ is fulfilled.
2. What is $\mathrm{P}(X + Y \leq 2)$?
3. Give the marginal probability mass functions of X

probability mass function Example

Consider the probability mass function, $f_{XY}(x,y)$, for the discrete random variables X and Y:

$$f_{XY}:\Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to \frac{xy-1}{6}$$

where $\Omega = \{1,2\} \times \{1,2\}$.

1 Show that $\sum\limits_i \sum\limits_j f_{XY}(x_i, y_j) = 1$ is fulfilled.

Denote $x_1 = 1, x_2 = 2, y_1 = 1$ and $y_2 = 2$. We then have $f_{XY}(1,1) = 0$, $f_{XY}(2,1) = f_{XY}(1,2) = \frac{1}{6}$ and $f_{XY}(2,2) = \frac{4}{6}$. Hence

$$\sum_{i=1}^{2} \sum_{j=1}^{2} f_{XY}(x_i, y_j) = 0 + \frac{1}{6} + \frac{1}{6} + \frac{4}{6} = 1$$

probability mass function Example

Consider the probability mass function, $f_{XY}(x,y)$, for the discrete random variables X and Y:

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to \frac{xy-1}{6}$$

where $\Omega = \{1,2\} \times \{1,2\}$.

2 What is $\mathrm{P}(X+Y \leq 2)$?

The set of values fulfilling $X+Y \leq 2$ is given by $\{(1,1)\}$. Hence,

$$\mathrm{P}(X+Y \leq 2) = \mathrm{P}(X+Y = 2) = f_{XY}(1,1) = 0$$

probability mass function Example

Consider the probability mass function, $f_{XY}(x,y)$, for the discrete random variables X and Y:

$$f_{XY}:\Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to \frac{xy-1}{6}$$

where $\Omega = \{1,2\} \times \{1,2\}$.

3 Give the marginal probability mass functions of X

$$f_X(1) = \sum_j f_{XY}(1,y_j) = f_{XY}(1,1) + f_{XY}(1,2) = 0 + \tfrac{1}{6} = \tfrac{1}{6}$$
$$f_X(2) = \sum_j f_{XY}(2,y_j) = f_{XY}(2,1) + f_{XY}(2,2) = \tfrac{1}{6} + \tfrac{4}{6} = \tfrac{5}{6}$$

Continuous Random Variables pdf (Recap)

- For a continuous rv X with CDF $F_X(x)$, its derivative

$$f_X(x) \equiv \frac{\partial F_X(x)}{\partial x}$$

is called the probability density function (pdf).

- From above's equation we conclude

$$F_X(x) \equiv \int_{-\infty}^{x} f_X(u) \, \mathrm{d}u \, .$$

⇝ The pdf is a function for which the area under the curve corresponding to any interval is equal to the probability that X will take on a value in that interval

Continuous Random Variables joint pdf

- For continuous rvs X and Y with CDF $F_{XY}(x,y)$, its derivative

$$f_{XY}(x,y) \equiv \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

  is called the (joint) probability density function (pdf).

- From above's equation we conclude

$$F_{XY}(x,y) \equiv \mathrm{P}(X \leq x \cap Y \leq y) \equiv \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(u,v) \; \mathrm{d}u \; \mathrm{d}v \; .$$

- Moreover, it holds that

$$\mathrm{P}(x_1 < X \leq x_2 \cap y_1 < Y \leq y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f_{XY}(x,y) \; \mathrm{d}x \; \mathrm{d}y \; .$$

Continuous Random Variables pdf Properties

- $f_{XY}(x,y) \geq 0 \forall x,y$ (since 'probability under curve')

  (*necessary* condition)

- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)\, \mathrm{d}x\, \mathrm{d}y = 1$ (since $\lim_{x,y \to \infty} F_{XY}(x,y) = 1$)

  (*necessary* condition)

- $\int_{-\infty}^{\infty} f_{XY}(x,y)\, \mathrm{d}y = f_X(x)$ (marginal pdf of X)

- $\int_{-\infty}^{\infty} f_{XY}(x,y)\, \mathrm{d}x = f_Y(y)$ (marginal pdf of Y)

DHBW Duale Hochschule Baden-Württemberg Mannheim

Recap: Expected Value (Lecture 2)

Let $g(X)$ be a real valued function of a rv X.

- Let X be a discrete rv with realisations x_i. Then

$$\mathrm{E}(g(X)) \equiv \sum_i g(x_i) \cdot f_X(x_i) \equiv \sum_i g(x_i) \cdot \mathrm{P}(X = x_i) \; .$$

- Let X be a continuous rv. Then

$$\mathrm{E}(g(X)) \equiv \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \, \mathrm{d}x \; .$$

Expected Value for Bivariate Distributions

Let $g : \mathbb{R}^2 \to \mathbb{R}, \begin{pmatrix} X \\ Y \end{pmatrix} \mapsto g(X,Y)$ be a real valued function of two rvs X and Y.

- Let X and Y be discrete with realisations x_i and y_j. Then

$$\mathrm{E}(g(X,Y)) \equiv \sum_i \sum_j g(x_i, y_j) \cdot f_{XY}(x_i, y_j) \; .$$

- Let X and Y be continuous. Then

$$\mathrm{E}(g(X,Y)) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \cdot f_{XY}(x,y) \, \mathrm{d}x \, \mathrm{d}y \; .$$

$$\rightsquigarrow E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot f_{XY}(x,y) \, \mathrm{d}x \, \mathrm{d}y = \int_{-\infty}^{\infty} x \cdot f_X(x) \, \mathrm{d}x$$

$$\rightsquigarrow E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \cdot f_{XY}(x,y) \, \mathrm{d}x \, \mathrm{d}y = \int_{-\infty}^{\infty} y \cdot f_Y(y) \, \mathrm{d}y$$

Expected Value for Bivariate Distributions (Example)

Let X and Y be continuous. How to calculate $\mathrm{E}(2X + X^2 - Y)$?

1. Recap: $\mathrm{E}(g(X,Y)) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \cdot f_{XY}(x,y) \, \mathrm{d}x \, \mathrm{d}y$.

2. Define $g : \mathbb{R}^2 \to \mathbb{R}, \begin{pmatrix} X \\ Y \end{pmatrix} \mapsto 2X + X^2 - Y$

3. Then $\mathrm{E}(2X + X^2 - Y) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (2x + x^2 - y) \cdot f_{XY}(x,y) \, \mathrm{d}x \, \mathrm{d}y$.

Bivariate Gaussian Distribution (Example)

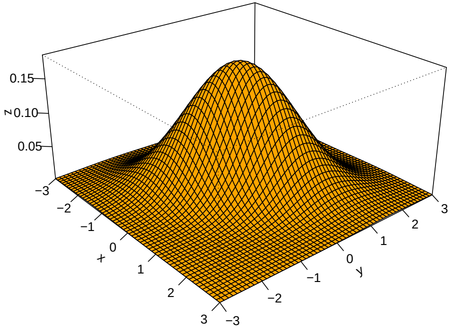

Figure: pdf for a bivariate Gaussian distribution

Table of Contents

1 Continuous Probability Distributions

2 Bivariate Random Variables

3 Bivariate Distribution: Independence, (Conditional) Moments

Table of Contents

1 Continuous Probability Distributions

2 Bivariate Random Variables

3 Bivariate Distribution: Independence, (Conditional) Moments

Recap: Conditional Probability (Lecture 1)

- Given two events A and B, the conditional probability of A given B is defined as

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \quad , \mathrm{P}(B) > 0 \ .$$

- Equivalently, it holds that

$$\mathrm{P}(A \cap B) = \mathrm{P}(A|B) \cdot \mathrm{P}(B) \ .$$

Conditional Distribution

- The conditional CDF of a rv X given that another rv Y has taken a value y, is defined by

$$F_{XY}(x|y) \equiv \mathrm{P}(X \leq x|Y = y) \ .$$

- For *discrete* X and Y, the conditional pmf of X given $Y = y$ is

$$f_{XY}(x|y) \equiv \mathrm{P}(X = x|Y = y) \equiv \frac{\mathrm{P}(X = x \cap Y = y)}{\mathrm{P}(Y = y)} \equiv \frac{f_{XY}(x,y)}{f_Y(y)} \ ,$$

if $\mathrm{P}(Y = y) = f_Y(y) > 0$.

- For *continuous* X and Y, the conditional pdf of X given $Y = y$ is

$$f_{XY}(x|y) \equiv \frac{\partial F_{XY}(x|y)}{\partial x} \equiv \frac{f_{XY}(x,y)}{f_Y(y)} \ ,$$

if $f_Y(y) > 0$.

Practical Properties

- When X and Y are discrete, then

$$F_{XY}(x|y) = \sum_{i}^{i:x_i \le x} f_{XY}(x_i|y) \, .$$

- When X and Y are continuous, then

$$F_{XY}(x|y) = \int_{-\infty}^{x} f_{XY}(u|y) \, \mathrm{d}u \, .$$

Recap: Independence (Lecture 1)

- Two events A and B are said to be independent, if

$$P(A|B) = P(A) \quad \Big(\text{or } P(B|A) = P(B)\Big) \ .$$

- Equivalently, it holds that

$$P(A \cap B) = P(A) \cdot P(B) \ .$$

- Where does this come from?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A) \Leftrightarrow P(A \cap B) = P(A) \cdot P(B) \ .$$

Independent Random Variables

- Two rvs X and Y are said to be independent, if $\forall x, y$

$$F_{XY}(x,y) = F_X(x)F_Y(y) \Leftrightarrow F_{XY}(x|y) = F_X(x) \Leftrightarrow F_{XY}(y|x) = F_Y(y) \, .$$

- Two *discrete* rvs X and Y are said to be independent, if $\forall x, y$

$$f_{XY}(x,y) = f_X(x)f_Y(y) \Leftrightarrow f_{XY}(x|y) = f_X(x) \Leftrightarrow f_{XY}(y|x) = f_Y(y) \, .$$

- Two *continuous* rvs X and Y are said to be independent, if $\forall x, y$

$$f_{XY}(x,y) = f_X(x)f_Y(y) \Leftrightarrow f_{XY}(x|y) = f_X(x) \Leftrightarrow f_{XY}(y|x) = f_Y(y) \, .$$

- Notation: If X and Y are independent, we write $X \perp\!\!\!\perp Y$.

DHBW Duale Hochschule Baden-Württemberg Mannheim

Independent Random Variables: Function of Random Variables

- Consider two random variables X and Y, and two continuous functions $g, h : \mathbb{R} \to \mathbb{R}$.

$$\text{If } X \perp\!\!\!\perp Y \;\Rightarrow\; g(X) \perp\!\!\!\perp h(Y)$$

- Example: Consider $X \perp\!\!\!\perp Y$. Then
  - $2X \perp\!\!\!\perp -Y \qquad \Big(\, g : x \mapsto 2x \text{ and } f : y \mapsto -y \Big)$
  - $X^2 \perp\!\!\!\perp Y + 4 \qquad \Big(\, g : x \mapsto x^2 \text{ and } f : y \mapsto y + 4 \Big)$

Extension: Function of independent Random Vectors

- Consider two independent random vectors

$$X = \Big(X_1, \ldots, X_m\Big)' \text{ and } Y = \Big(Y_1, \ldots, Y_m\Big)' ,$$

  and two continuous functions

$$g : \mathbb{R}^m \to \mathbb{R}^u \text{ and } h : \mathbb{R}^n \to \mathbb{R}^v .$$

  It then holds that

$$g(X) \perp\!\!\!\perp h(Y) .$$

- Example: Consider $(X_1, X_2) \perp\!\!\!\perp (Y_1, Y_2)$. Then
  - $X_1 \perp\!\!\!\perp Y_2 \quad \Big(g : (x_1, x_2)' \mapsto x_1 \text{ and } h : (y_1, y_2)' \mapsto y_1 \Big)$
  - $X_1 \perp\!\!\!\perp Y_2 \quad \Big(g : (x_1, x_2)' \mapsto x_1 \text{ and } h : (y_1, y_2)' \mapsto y_2 \Big)$
  - $X_1 \perp\!\!\!\perp Y_1 + Y_2 \quad \Big(g : (x_1, x_2)' \mapsto x_1 \text{ and } h : (y_1, y_2)' \mapsto y_1 + y_2 \Big)$

Recap: Expected Value (Lecture 2)

Let $g(X)$ be a real valued function of a rv X.

- Let X be a discrete rv with realisations x_i. Then

$$\mathrm{E}(g(X)) \equiv \sum_i g(x_i) \cdot f_X(x_i) \equiv \sum_i g(x_i) \cdot \mathrm{P}(X = x_i) \,.$$

- Let X be a continuous rv. Then

$$\mathrm{E}(g(X)) \equiv \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \, \mathrm{d}x \,.$$

Conditional Expected Value

Let $g(X)$ be a real valued function of a rv X.

- For *discrete* Y and X with realisation x_i, the conditional expected value (expectation/mean) of X given $Y = y$ is

$$\mathrm{E}(g(X)|Y = y) \equiv \sum_i g(x_i) \cdot f_{X\,Y}(x_i|y) \equiv \sum_i g(x_i) \cdot \mathrm{P}(X = x_i|Y = y) \ .$$

- For *continuous* Y and X with realisation x_i, the conditional expected value (expectation/mean) of X given $Y = y$ is

$$\mathrm{E}(g(X)|Y = y) \equiv \int_{-\infty}^{\infty} g(x) \cdot f_{X\,Y}(x_i|y) \ \mathrm{d}x \ .$$

Conditional Expected Value: Properties

Let X and Y be two random variables. It holds that

- Law of Iterated Expectations (LIE): $\mathrm{E}(X) = \mathrm{E}(\mathrm{E}(X|Y))$
- $\mathrm{E}(f(X)Y|X) = f(X) \cdot \mathrm{E}(Y|X)$ (i.e. $\mathrm{E}(XY|X) = X \cdot \mathrm{E}(Y|X)$ and $\mathrm{E}(X|X) = X$)

Let X and Y be two random variables with $X \perp\!\!\!\perp Y$. It holds that

- $\mathrm{E}(XY) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$.
- $\mathrm{E}(X|Y) = \mathrm{E}(X)$.

DHBW Duale Hochschule Baden-Württemberg Mannheim

probability mass function Example (continued)

Consider the probability mass function, $f_{XY}(x,y)$, for the discrete random variables X and Y:

$$f_{XY}:\Omega\to\mathbb{R},\ \begin{pmatrix}x\\y\end{pmatrix}\to\frac{xy-1}{6}$$

where $\Omega=\{0,1\}\times\{0,1\}$.

4 What is $\mathrm{E}(X|Y=1)$?

5 What is $\mathrm{E}(Y|X=2)$?

6 Is $X \perp\!\!\!\perp Y$?

probability mass function Example (continued)

Consider the probability mass function, $f_{XY}(x, y)$, for the discrete random variables X and Y:

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to \frac{xy - 1}{6}$$

where $\Omega = \{1, 2\} \times \{1, 2\}$.

4 What is $\mathrm{E}(X|Y = 1)$?

$$\begin{aligned} &\mathrm{E}(X|Y = 1) \\ \equiv\ & \sum_{i=1}^{2} x_i f_{XY}(x_i|1) \\ \equiv\ & \sum_{i=1}^{2} x_i \frac{f_{XY}(x_i, 1)}{f_Y(1)} \\ =\ & 1 \cdot \frac{f_{XY}(1,1)}{f_Y(1)} + 2 \cdot \frac{f_{XY}(2,1)}{f_Y(1)} \\ =\ & 1 \cdot \frac{0}{1/6} + 2 \cdot \frac{1/6}{1/6} \\ =\ & 2 \end{aligned}$$

DHBW
Duale Hochschule
Baden-Württemberg
Mannheim

probability mass function Example (continued)

Consider the probability mass function, $f_{XY}(x,y)$, for the discrete random variables X and Y:

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to \frac{xy-1}{6}$$

where $\Omega = \{1,2\} \times \{1,2\}$.

5 What is $\mathrm{E}(Y|X=2)$?

$$\begin{aligned}
&\mathrm{E}(Y|X=2) \\
&= \sum_{i=1}^{2} y_i f_{Y\,X}(y_i|2) \\
&= \sum_{i=1}^{2} y_i \frac{f_{Y\,X}(y_i,2)}{f_X(2)} \\
&= 1 \cdot \frac{f_{Y\,X}(1,2)}{f_X(2)} + 2 \cdot \frac{f_{Y\,X}(2,2)}{f_X(2)} \\
&= 1 \cdot \frac{1/6}{5/6} + 2 \cdot \frac{4/6}{5/6} \\
&= \frac{9}{5}
\end{aligned}$$

Consider the probability mass function, $f_{XY}(x,y)$, for the discrete random variables X and Y:

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to \frac{xy-1}{6}$$

where $\Omega = \{1,2\} \times \{1,2\}$.

6 Is $X \perp\!\!\!\perp Y$?

- $f_{XY}(1,1) = 0 \neq \frac{1}{6} \cdot \frac{1}{6} = f_X(1) f_Y(1) \rightsquigarrow X \not\perp\!\!\!\perp Y$
- $f_{XY}(1,2) = \frac{1}{6} \neq \frac{1}{6} \cdot \frac{5}{6} = f_X(1) f_Y(2)$
- $f_{XY}(2,1) = \frac{1}{6} \neq \frac{5}{6} \cdot \frac{1}{6} = f_X(2) f_Y(1)$
- $f_{XY}(2,2) = \frac{4}{6} \neq \frac{5}{6} \cdot \frac{5}{6} = f_X(2) f_Y(2)$

probability density function Example

Consider the probability density function, $f_{XY}(x,y)$, for the continuous random variables X and Y:

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy$$

where $\Omega = [0,1] \times [0,1]$.

1. Show that $f_{XY}(x,y)$ is a valid pdf.
2. What is $\mathrm{P}(Y < X)$?
3. Give the marginal pdfs of X and Y.
4. Is $X \perp\!\!\!\perp Y$?
5. Calculate $\mathrm{E}(X)$ and $\mathrm{E}(Y)$
6. Calculate $\mathrm{E}(2X + \frac{1}{2}Y)$

probability density function Example

$$f_{XY}:\Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy \ , \quad \Omega = [0,1] \times [0,1]$$

1 Show that $f_{XY}(x,y)$ is a valid pdf.

(i) $f_{XY} \geq 0 \ \forall x, y \in [0,1]$ ✓

(ii) $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f_{XY} \ \mathrm{d}x \ \mathrm{d}y = 1$:

$$\begin{aligned}
& \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f_{XY} \ \mathrm{d}x \ \mathrm{d}y \\
= \quad & \int\limits_{0}^{1} \int\limits_{0}^{1} f_{XY} \ \mathrm{d}x \ \mathrm{d}y \\
= \quad & \int\limits_{0}^{1} \int\limits_{0}^{1} 4xy \ \mathrm{d}x \ \mathrm{d}y \\
= \quad & \int\limits_{0}^{1} \left[2x^2 y\right]_{x=0}^{x=1} \ \mathrm{d}y \\
= \quad & \int\limits_{0}^{1} 2y \ \mathrm{d}y \\
= \quad & \left[y^2\right]_0^1 = 1 \ \checkmark
\end{aligned}$$

DHBW Duale Hochschule Baden-Württemberg Mannheim

probability density function Example

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy \ , \quad \Omega = [0,1] \times [0,1]$$

2 What is $P(Y < X)$?

We consider the plane where $\{0 \leq x \leq 1\}$ and $\{0 \leq y \leq x\}$:

$$\begin{aligned}
\rightsquigarrow \quad & \int_0^x \int_0^1 f_{XY} \ \mathrm{d}x \ \mathrm{d}y \\
= \quad & \int_0^1 \int_0^x f_{XY} \ \mathrm{d}y \ \mathrm{d}x \\
= \quad & \int_0^1 \int_0^x 4xy \ \mathrm{d}y \ \mathrm{d}x \\
= \quad & \int_0^1 \left[2xy^2\right]_{y=0}^{y=x} \ \mathrm{d}x \\
= \quad & \int_0^1 2x^3 \ \mathrm{d}x \\
= \quad & \left[\tfrac{1}{2}x^4\right]_0^1 = \tfrac{1}{2}
\end{aligned}$$

probability density function Example

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy \ , \quad \Omega = [0,1] \times [0,1]$$

3 Give the marginal pdfs of X and Y.

$$f_X = \int\limits_{-\infty}^{\infty} f_{XY} \ \mathrm{d}y = \int\limits_{0}^{1} 4xy \ \mathrm{d}y = \left[2xy^2\right]_{y=0}^{y=1} = 2x$$

$$\rightsquigarrow f_X : [0,1] \to \mathbb{R}, \ x \mapsto 2x \ .$$

$$f_Y = \int\limits_{-\infty}^{\infty} f_{XY} \ \mathrm{d}x = \int\limits_{0}^{1} 4xy \ \mathrm{d}x = \left[2x^2y\right]_{x=0}^{x=1} = 2y$$

$$\rightsquigarrow f_Y : [0,1] \to \mathbb{R}, \ x \mapsto 2y \ .$$

probability density function Example

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy \;, \quad \Omega = [0,1] \times [0,1]$$

4 Is $X \perp\!\!\!\perp Y$?

$$f_X \cdot f_Y = 2x \cdot 2y = 4xy = f_{XY} \quad \Big(\text{for } x, y \in [0,1] \Big)$$

$$\rightsquigarrow X \perp\!\!\!\perp Y \;\checkmark$$

DHBW Duale Hochschule Baden-Württemberg Mannheim

$$f_{XY}: \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy\ , \quad \Omega = [0,1] \times [0,1]$$

5 Calculate $\mathrm{E}(X)$ and $\mathrm{E}(Y)$

$$\mathrm{E}X = \int\limits_{-\infty}^{\infty} x f_X \ \mathrm{d}x = \int\limits_0^1 x \cdot 2x \ \mathrm{d}x = \int\limits_0^1 2x^2 \ \mathrm{d}x = \left[\tfrac{2}{3}x^3\right]_0^1 = \tfrac{2}{3}$$

$$\mathrm{E}Y = \ldots = \tfrac{2}{3}$$

probability density function Example

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy \ , \quad \Omega = [0,1] \times [0,1]$$

6 Calculate $\mathrm{E}(2X + \frac{1}{2}Y)$

$$\begin{aligned}
& \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (2x + \tfrac{1}{2}y) f_{XY} \ \mathrm{d}x \ \mathrm{d}y \\
= \quad & \int_0^1 \int_0^1 (2x + \tfrac{1}{2}y) 4xy \ \mathrm{d}x \ \mathrm{d}y \\
= \quad & \int_0^1 \int_0^1 8x^2 y + 2xy^2 \ \mathrm{d}x \ \mathrm{d}y \\
= \quad & \int_0^1 \left[\tfrac{8}{3} x^3 y + x^2 y^2 \right]_{x=0}^{x=1} \ \mathrm{d}y \\
= \quad & \int_0^1 \tfrac{8}{3} y + y^2 \ \mathrm{d}y \\
= \quad & \left[\tfrac{4}{3} y^2 + \tfrac{1}{3} y^3 \right]_0^1 \\
= \quad & \tfrac{4}{3} + \tfrac{1}{3} = \tfrac{5}{3}
\end{aligned}$$

probability density function Example

$$f_{XY} : \Omega \to \mathbb{R}, \begin{pmatrix} x \\ y \end{pmatrix} \to 4xy \ , \quad \Omega = [0,1] \times [0,1]$$

5 Calculate E(X) and E(Y): Remark

We can calculate EX and EY without using the respective marginal densities:

$$\mathrm{E}X = \mathrm{E}(X + 0 \cdot Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + 0 \cdot y) f_{XY} \ \mathrm{d}x \ \mathrm{d}y = \frac{2}{3}$$